

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2001-155467
(P2001-155467A)

(43) 公開日 平成13年6月8日 (2001.6.8)

(51) Int.Cl.

識別記号

F I

テマコード* (参考)

G 1 1 B 27/031
G 0 6 F 17/22
G 1 0 L 15/00

G 1 1 B 27/02
G 0 6 F 15/20
G 1 0 L 3/00

H 5 B 0 0 9
5 0 3 5 D 0 1 5
5 5 1 B 9 A 0 0 1
5 5 1 G

審査請求 有 請求項の数 6 O L (全 15 頁)

(21) 出願番号 特願平11-337041

(22) 出願日 平成11年11月29日 (1999. 11. 29)

(71) 出願人 390024350

株式会社ジャストシステム
徳島県徳島市沖浜東 3-46

(72) 発明者 浮川 初子

徳島県徳島市沖浜東 3丁目46番地 株式会
社ジャストシステム内

(72) 発明者 森本 寛

徳島県徳島市沖浜東 3丁目46番地 株式会
社ジャストシステム内

(74) 代理人 100096655

弁理士 川井 隆 (外1名)

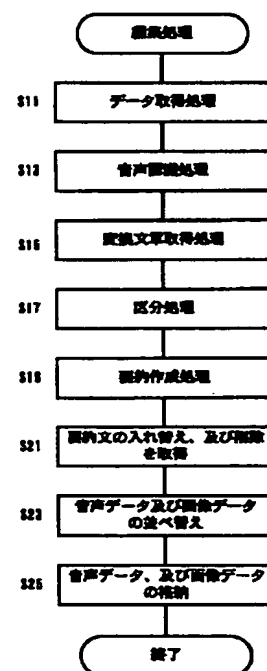
最終頁に続く

(54) 【発明の名称】 編集処理装置、及び編集処理プログラムが記憶された記憶媒体

(57) 【要約】

【課題】 音声データの編集を容易に短時間で行うことのできる編集処理装置を提供すること。

【解決手段】 編集処理装置は、音声認識による音声認識結果を音声データ、及びその録音時刻とともに取得する (S11、S13)。そして音声認識結果に基づいて文章を作成し (S15)、段落毎に対応する録音時間を取得するとともに段落毎の要約を作成する (S17、S19)。そして要約を表示装置に表示し、段落単位での入れ替えや削除等を取得する (S21)。続いて、段落に対応する録音時刻から、入れ替えされたり削除された段落に対応する音声データを把握し、音声データも表示された文章の入れ替えや削除に従って入れ替え及び削除を行う (S23)。



1

【特許請求の範囲】

【請求項1】 音声認識手段による音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得手段と、
前記認識結果取得手段で取得した前記音声認識結果に基づいて認識文章を作成する認識文章作成手段と、
前記文章作成手段で作成された前記認識文章と前記認識結果取得手段で取得した前記音声データとの対応を前記認識文章の所定単位毎に取得する対応取得手段と、
前記文章作成手段で取得された前記認識文章を前記所定単位で編集する文章編集手段と、
前記文書編集手段の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集手段とを備えることを特徴とする編集処理装置。

【請求項2】 前記所定単位は、音素、単音節、単語、形態素、文節、及び文、のうちのいずれかであることを特徴とする請求項1に記載の編集処理装置。

【請求項3】 前記文章作成手段は、前記認識結果取得手段で取得した音声認識結果を文章に変換する変換手段と、前記変換手段により取得された変換文章を、複数の段落の段落に区分する区分手段とを備え、
前記所定単位は、前記区分手段で区分された前記段落であることを特徴とする請求項1に記載の編集処理装置。

【請求項4】 前記文章作成手段は、前記認識結果取得手段で取得した音声認識結果を文章に変換する変換手段と、前記変換手段により取得された変換文章を、複数の段落の段落に区分する区分手段と、前記区分手段で区分された前記段落それぞれを要約する要約手段とを備え、
前記要約手段により取得された各段落の前記要約からなる文章を認識文書として取得し、
前記対応取得手段は、前記段落を前記所定単位として、前記認識文章と前記音声データとの対応を取得することを特徴とする請求項1に記載の編集処理装置。

【請求項5】 前記認識結果取得手段は、画像データに付加された音声データを前記画像データとともに取得し、
前記対応取得手段は、前記文章作成手段で作成された前記認識文章と、前記音声データ及び前記画像データとの所定単位毎の対応を取得し、
前記データ編集手段は、前記文書編集手段の編集結果に従って、前記対応に基づいて、前記音声データとともに前記画像データを編集することを特徴とする請求項1から請求項4のうちのいずれか1の請求項に記載の編集処理装置。

【請求項6】 音声認識手段による音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得機能と、
前記認識結果取得機能で取得した前記音声認識結果に基づいて認識文章を作成する文章作成機能と、
前記文章作成機能で作成された前記認識文章と前記認識

2

結果取得機能で取得した前記音声データと対応を前記認識文章の所定単位毎に取得する対応取得機能と、
前記文章作成機能で取得された前記認識文章を前記所定単位で編集する文章編集機能と、
前記文書編集機能の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集機能とをコンピュータに実現させるためのコンピュータ読み取り可能な編集処理プログラムが記憶された記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、編集処理装置、及び編集処理プログラムが記憶された記憶媒体に関し、更に詳細には、音声データを容易に編集可能な編集処理装置、及び編集処理プログラムが記憶された記憶媒体に関する。

【0002】

【従来の技術】 従来、音声データを編集処理装置、及び編集処理プログラムが記憶された記憶媒体による編集処理によって編集する場合には、実際に音声を再生することによって編集したい音声データの位置を検出したり、音声データとともにタイムレコードを記録しておき、タイムレコードを目安にして編集したい位置を検出している。そして、音声やタイムレコードを目安にして、音声データの波形等を扱って手動で音声データを編集している。

【0003】

【発明が解決しようとする課題】 しかし、一連の音声データを区切って編集した位置を指定するためには、微妙なタイミングでの操作が必要であり、熟練を必要とする問題点がある。また、タイムレコード等の、音声以外の目安を使用する場合には、音声データを取り扱うため、直感的な取り扱いができない問題点がある。

【0004】 尚、音声データを伴う映像データを編集する編集処理においては、音声データとを文字データに変換して表示し、文字データを利用して映像データを編集する技術が提案されている（特開平9-130736号公報）。この編集処理においては、音声データを変換した文字データによって編集箇所を検出する。そのため、音声の再生やタイムレコードを目安にするよりも容易に編集位置を探すことができる。音声データとしては、撮像時の音声のデータや、撮像時に編集用に入力した音声のデータが使用される。しかし、映画やニュースでは、撮像時の音声のデータは膨大となるため、その音声データを文字データに変換しても、文字データの量が多く、編集箇所を検出するのは時間がかかる場合がある。また、編集用の音声のデータを使用する場合には、その音声を入力するための手間がかかる問題点がある。更に、この技術では、編集位置の検索以外については、従来と同様に音声データや映像データを手動で取り扱うことで行われ、手間が改善されていない問題点がある。この技

術では、文字データを編集することはできず、文字データの編集によって音声データが編集される訳ではない。

【0005】本発明は、上述のような課題を解決するためになされたもので、音声データの編集を容易に短時間で行うことのできる編集処理装置、及び編集処理プログラムが記憶された記憶媒体を提供することを目的とする。

【0006】

【課題を解決するための手段】上記目的を達成するために、請求項1に記載の発明は、音声認識手段による音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得手段と、前記認識結果取得手段で取得した前記音声認識結果に基づいて認識文章を作成する認識文章作成手段と、前記文章作成手段で作成された前記認識文章と前記認識結果取得手段で取得した前記音声データとの対応を前記認識文章の所定単位毎に取得する対応取得手段と、前記文章作成手段で取得された前記認識文章を前記所定単位で編集する文章編集手段と、前記文書編集手段の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集手段とを備える編集処理装置を提供することにより前記目的を達成するものである。

【0007】上述の音声認識結果は、音素、単音節、単語、形態素、等、通常の音声認識手段（音声認識装置）によって、音声波形等の音声データを音声認識辞書と照合し取得される音声認識結果、またはこれらを仮名漢字変換したものである。この仮名漢字変換は、自然言語処理に基づいた解析を加えて変換したものとすることができる。また、音声認識結果は、外部の音声認識装置による音声認識結果であっても、本編集処理装置に配設された音声認識部による結果であってもよい。外部の音声認識装置による音声認識結果は、I/Oインターフェースを介して取得されたものであっても、フロッピーディスクやフラッシュメモリ等の各種記憶媒体や、インターネット及びパソコン通信等の公衆回線や専用回線を介して取得されたものであってもよい。上記文章作成手段は、前記音声認識結果によって取得された音素、単音節、単語、形態素を仮名漢字変換して認識文章とするものとする。作成する文章は、1つの文であっても、複数の文から構成されたものであってもよい。また、音声認識装置への音声入力時に、「次の段落」等の音声によって段落に区分されて複数の段落に構成されたものであってもよい。また、音声認識結果として仮名漢字変換語のデータを取得した場合には、このまま文章としたり、取得したデータに独自の自然言語処理等による分析を加えて再度仮名漢字変換し直した文章とするものであってもよい。上記対応取得手段は、音声認識結果を取得した先の音声認識手段から、音声認識結果と音声データとの対応を取得する。この対応は、音声認識結果と、音声データとのそれぞれについて、所定時を基準とし

た時間を介しての対応とすることができる。例えば、音声認識結果の所定単位である「わたくしは」が、認識開始後0.7秒から1.1秒までの間であった場合には、音声データのうち認識開始後0.7から1.1秒までの部分が、認識結果である「わたくしは」に対応しているとする。対応の前記所定単位は、音声認識時の単位と同じ細かさでもよいが、音声認識の単位のまとまりであってもよい。たとえば、単語毎に音声認識結果が取得された場合であっても、対応は、認識された単語が集合して形成された文節や文、段落毎に、音声データとの対応が取得されるようにしてもよい。

【0008】請求項2に記載の発明は、請求項1に記載の編集処理装置において、前記所定単位は、音素、単音節、単語、形態素、文節、及び文のうちのいずれかである編集処理装置を提供することにより前記目的を達成するものである。

【0009】請求項3に記載の発明は、前記文章作成手段は、前記認識結果取得手段で取得した音声認識結果を文章に変換する変換手段と、前記変換手段により取得された変換文章を、複数の段落の段落に区分する区分手段とを備え、前記所定単位は、前記区分手段で区分された前記段落である請求項1に記載の編集処理装置を提供することにより前記目的を達成するものである。前記区分手段は、例えば、変換手段で取得された変換文章中に頻出する自立語等を重要語（キーワード）として複数調べ出し（重要語取得手段）、最も出現頻度の高い重要語が切り替わる位置を段落の分け目として段落区分を取得する（区分決定手段）ものとする。また、前記区分手段は、所定の数の文のまとまり毎に1または複数の重要語を調べ出し、所定のまとまり毎に、そのまとまりの中でもっとも重要語の出現頻度の高い重要文（要約）を選び出し、重要語どうしの関連等から各重要文どうしの類似度を算出し、類似度が所定以上の重要文を含むまとまりどうしは同一の段落とし、類似度が低いまとまりどうしの間で段落が変わるとして区分するものとする。また、前記区分手段は、「次に」「さて、」等の所定の1または複数の接続詞その他の単語（特定語から後は新しい段落としたり、「以上」「終わり」「お伝えしました」等の特定の1または複数の単語（特定語）の後は、新しい段落とする等、特定語を検出し（特定語検出手段）、この特定語に基づいて文書を段落に区分するものとする。更に、前記区分手段は、音声認識時に、「新しい段落」等の、段落区分のための音声が入力されたものを認識して段落が形成されている場合には、この段落をそのまま認識文章の段落としてもよい。

【0010】請求項4に記載の発明は、前記文章作成手段は、前記認識結果取得手段で取得した音声認識結果を文章に変換する変換手段と、前記変換手段により取得された変換文章を、複数の段落の段落に区分する区分手段

5

と、前記区分手段で区分された前記段落それぞれを要約する要約手段とを備え、前記要約手段により取得された各段落の前記要約からなる文章を認識文書として取得し、前記対応取得手段は、前記段落を前記所定単位として、前記認識記文章と前記音声データとの対応を取得する請求項1に記載の編集処理装置を提供することにより前記目的を達成する。前記要約手段としては、変化文章中から重要語（キーワード）を検出し（重要語検出手段）、各段落中から重要語の出現頻度の大きな1または複数の文等を重要文として選択し（文選択手段）、これらの文をその文を含む段落順にならべて（文整列手段）要約とすることができる。

【0011】請求項5に記載の発明は、前記認識結果取得手段は、画像データに付加された音声データを前記画像データとともに取得し、前記対応取得手段は、前記文章作成手段で作成された前記認識文章と、前記音声データ及び前記画像データとの所定単位毎の対応を取得し、前記データ編集手段は、前記文書編集手段の編集結果に従って、前記対応に基づいて、前記音声データとともに前記画像データを編集する請求項1から請求項4のうちのいずれか1の請求項に記載の編集処理装置を提供することにより前記目的を達成するものである。

【0012】請求項6に記載の発明は、音声認識手段による音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得機能と、前記認識結果取得機能で取得した前記音声認識結果に基づいて認識文章を作成する認識文章作成機能と、前記文章作成機能で作成された前記認識文章と前記認識結果取得機能で取得した前記音声データとの対応を前記認識文章の所定単位毎に取得する対応取得機能と、前記文章作成機能で取得された前記認識文章を前記所定単位で編集する文章編集機能と、前記文書編集機能の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集機能と、をコンピュータに実現させるためのコンピュータ読み取り可能な編集処理プログラムが記憶されたことを特徴とする記憶媒体を提供することにより前記目的を達成するものである。また、前記認識結果取得機能と、前記認識文章作成機能と、前記対応取得機能と、文章編集機能と、前記データ編集機能と、を実現するための編集処理プログラム、編集処理プログラム伝送媒体、編集処理プログラム搬送波、編集処理プログラム信号、またはプログラム製品としてもよい。ここで、プログラム製品には、編集処理プログラムによる前記各機能を実現する記憶媒体、サーバシステムコンピュータ、及びコンピュータシステム等を含む。

【0013】請求項6に記載の発明は、音声認識手段による音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得機能と、前記認識結果取得機能で取得した前記音声認識結果に基づいて認識文章を作成する認識文章作成機能と、前記文章作成機能で作成

6

された前記認識文章と前記認識結果取得機能で取得した前記音声データとの対応を前記認識文章の所定単位毎に取得する対応取得機能と、前記文章作成機能で取得された前記認識文章を前記所定単位で編集する文章編集機能と、前記文書編集機能の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集機能と、をコンピュータに実現させるためのコンピュータ読み取り可能な編集処理プログラムであって、且つ、前記所定単位は、音素、単音節、単語、形態素、文節、及び文、のうちのいずれかである編集処理プログラムが記憶される記憶媒体とすることができる。また、前記編集処理プログラム、編集処理プログラム伝送媒体、編集処理プログラム搬送波、編集処理プログラム信号、またはプログラム製品とすることができる。

【0014】請求項6に記載の発明は、音声認識手段による音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得機能と、前記認識結果取得機能で取得した前記音声認識結果に基づいて認識文章を作成する認識文章作成機能と、前記文章作成機能で作成された前記認識文章と前記認識結果取得機能で取得した前記音声データとの対応を前記認識文章の所定単位毎に取得する対応取得機能と、前記文章作成機能で取得された前記認識文章を前記所定単位で編集する文章編集機能と、前記文書編集機能の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集機能と、をコンピュータに実現させるためのコンピュータ読み取り可能な編集処理プログラムであって、且つ、前記文章作成機能は、前記認識結果取得機能で取得した音声認識結果を文章に変換する変換機能と、前記変換機能により取得された変換文章を、複数の段落の段落に区分する区分機能とを備え、前記所定単位は、前記区分機能で区分された前記段落である編集処理プログラムが記憶された記憶媒体とすることができる。また、前記編集処理プログラム、編集処理プログラム伝送媒体、編集処理プログラム搬送波、編集処理プログラム信号、またはプログラム製品とすることができる。前記編集処理プログラム、編集処理プログラム伝送媒体、編集処理プログラム搬送波、編集処理プログラム信号、またはプログラム製品は、音声入力手段を備えたコンピュータに、音声入力手段により入力された音声認識する音声認識機能を実現させるものとしてすることができる。

【0015】請求項6に記載の発明は、音声認識手段による音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得機能と、前記認識結果取得機能で取得した前記音声認識結果に基づいて認識文章を作成する認識文章作成機能と、前記文章作成機能で作成された前記認識文章と前記認識結果取得機能で取得した前記音声データとの対応を前記認識文章の所定単位毎に取得する対応取得機能と、前記文章作成機能で取得された前記認識文章を前記所定単位で編集する文章編集機能

7

と、前記文書編集機能の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集機能と、をコンピュータに実現させるためのコンピュータ読み取り可能な編集処理プログラムであって、且つ、前記文章作成機能は、前記認識結果取得機能で取得した音声認識結果を文章に変換する変換機能と、前記変換機能により取得された変換文章を、複数の段落の段落に区分する区分機能と、前記区分機能で区分された前記段落それぞれを要約する要約機能とを備え、前記要約機能により取得された各段落の前記要約からなる文章を認識文書として取得し、前記対応取得機能は、前記段落を前記所定単位として、前記認識記文章と前記音声データとの対応を取得する編集処理プログラムが記憶された記憶媒体とすることができる。また、前記編集処理プログラム、編集処理プログラム伝送媒体、編集処理プログラム搬送波、編集処理プログラム信号、またはプログラム製品とすることができる。

【0016】請求項6に記載の発明は、音声認識手段による音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得機能と、前記認識結果取得機能で取得した前記音声認識結果に基づいて認識文章を作成する認識文章作成機能と、前記文章作成機能で作成された前記認識文章と前記認識結果取得機能で取得した前記音声データとの対応を前記認識文章の所定単位毎に取得する対応取得機能と、前記文章作成機能で取得された前記認識文章を前記所定単位で編集する文章編集機能と、前記文書編集機能の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集機能と、をコンピュータに実現させるためのコンピュータ読み取り可能な編集処理プログラムであって、且つ、前記認識結果取得機能は、画像データに付加された音声データを前記画像データとともに取得し、前記対応取得機能は、前記文章作成機能で作成された前記認識文章と、前記音声データ及び前記画像データとの所定単位毎の対応を取得し、前記データ編集機能は、前記文書編集機能の編集結果に従って、前記対応に基づいて、前記音声データとともに前記画像データを編集する編集処理プログラムが記憶された記憶媒体とすることができる。また、前記編集処理プログラム、編集処理プログラム伝送媒体、編集処理プログラム搬送波、編集処理プログラム信号、またはプログラム製品とすることができる。

【0017】

【発明の実施の形態】以下、本発明の編集処理装置、及び編集処理プログラムが記憶された記憶媒体の好適な実施の形態について、図1から図10を参照して詳細に説明する。図1は、本発明の編集処理装置の第1の実施形態であり、本発明の編集処理プログラムが記憶された記憶媒体の第1の実施形態の該プログラムが読み取られた、コンピュータの構成を表したブロック図である。この図1に示すように、編集処理装置（コンピュータ）

8

は、装置全体を制御するための制御部11を備えている。この制御部11には、データバス等のバスライン21を介して、入力装置としてのキーボード12やマウス13、表示装置14、音声認識装置の一部としてのマイク15、記憶装置16、記憶媒体駆動装置17、通信制御装置18、入出力I/F19、及び音声出力装置20が接続されている。

【0018】制御部11は、CPU111、ROM112、RAM113を備えており、CPUがRAMをワーキングエリアとしてROMに格納されたプログラムを実行することによって、キーボードから入力された仮名文字列を漢字混り文に変換する仮名漢字変換処理、マイク15からの音声入力等によって音声データを取得する音声データ取得処理、音声データを音声認識辞書と比較対照して形態素分析を行って音声認識結果を取得する音声認識処理、音声認識処理や外部から入力された音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得処理、前記認識結果取得処理で取得した音声認識結果を文章に変換する変換処理、前記変換処理により取得された変換文章を、複数の段落に区分する区分処理、区分処理で区分された前記段落それぞれを要約する要約処理、要約処理により取得された各段落の前記要約からなる文章を認識文書として取得する認識文書取得処理、前記段落を前記所定単位として、前記認識記文章と前記音声データとの対応を取得する対応取得処理、要約を各段落毎に編集する文章編集処理、文書編集処理の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集処理、音声データや編集処理によって編集された音声データに基づいて音声出力装置20のスピーカから音声を出力する音声出力処理の各処理を行う。ROM112は、CPU111が各種制御や演算を行うための各種プログラムやデータが予め格納されたリードオンリーメモリである。

【0019】RAM113は、CPU111にワーキングメモリとして使用されるランダムアクセスメモリである。このRAM113には、本実施形態による編集処理を行うためのエリアとして、マイク15から入力された音声の音声データ、通信制御装置18や記憶媒体駆動装置17から入力された音声データが、音声認識開始時からの時刻とともに格納される音声データ格納エリア1131、音声データに形態素分析を行った形態素を解析して文章化した変換文章が形態素毎の音声認識開始時からの時刻とともに格納される変換文章格納エリア1132、変換文章の段落毎に抽出された重要文を、対応する各段落の変換文章の音声認識開始からの時間とともに格納する要約格納エリア1134、入出力I/F19を介して外部の撮像装置から入力された画像の画像データ、通信制御装置18や記憶媒体駆動装置17から前記音声データに対応する画像データが、時刻とともに格納される画像データ格納エリア1135、その他の各種エリアが確

保されるようになっている。

【0020】キーボード12は、かな文字を入力するためのかなキーやテンキー、各種機能を実行するための機能キー、カーソルキー、等の各種キーが配置されている。マウス13は、ポインティングデバイスであり、表示装置14に表示される文書中の一部を指定したり、表示装置14に表示されたアイコンやコマンドをクリック等によって選択し、対応する機能を指定する入力装置である。このキーボード12やマウス13は、編集手段の一部として機能し、キーボード12からの入力やマウス13によるポインティングによって、表示装置14に表示される重要文を並べ替えたり、削除する等の編集を行う。

【0021】表示装置14は、例えばCRTや液晶ディスプレイ等が使用される。この表示装置14には、外部からの音声データを文章化した変換文章の要約が表示され、要約文表示手段として機能する。また、表示装置14は、表示された要約を、要約文単位で、キーボードからの入力やマウス操作によって編集した編集要約を表示する編集要約表示手段としても機能する。

【0022】マイク15は、制御部11内の音声認識部（音声認識手段）において音声認識の対象となる音声が入力される音声入力手段として機能する。マイク15からの音声は、制御部11内の音声認識部により単語として認識される。そして、編集手段の一部として機能し、表示装置14に要約を表示させたり、要約文の入れ換えや削除、複写等の編集操作を行うことができるようになっている。このマイク15は、例えば音楽編集等の、本実施形態のコンピュータにおいて実行される他のプログラムにおいて必要な音声入力手段としても機能するものとすることもできる。また、入力音声を的確に収集するために指向性のある専用のマイクとしてもよい。更に、音声出力装置として図示しないスピーカ（イヤホン）を備え、このスピーカとマイク15とでハンズフリーユニットを形成させ、入力音声を確認しながら音声入力ができるようにしてもよい。

【0023】記憶装置16は、読み書き可能な記憶媒体と、その記憶媒体に対してプログラムやデータ等の各種情報を読み書きするための駆動装置で構成されている。この記憶装置16に使用される記憶媒体としては、主としてハードディスクが使用されるが、後述の記憶媒体駆動装置17で使用される各種記憶媒体のうちの読み書き可能な記憶媒体を使用するようにしてもよい。記憶装置16は、仮名漢字変換辞書格納部161、プログラム格納部162、文書格納部164、音声認識辞書格納部165、編集データ格納部166及び図示しないその他の格納部（例えば、この記憶装置16内に格納されているプログラムやデータ等をバックアップするための格納部）等を有している。

【0024】プログラム格納部162には、仮名漢字変

換辞書格納部に格納される仮名漢字変換辞書を使用して入力された仮名文字列を漢字混り文に変換する仮名漢字変換プログラム、マイク15からの音声入力により音声データを取得する音声データ取得処理プログラム、音声データを音声認識辞書と比較対照して形態素分析を行って音声認識結果を取得する音声認識処理プログラム、音声認識処理や外部から入力された音声認識結果を、認識の対象となった音声データとともに取得する認識結果取得処理プログラム、前記認識結果取得処理で取得した音声認識結果を文章に変換する変換処理プログラム、前記変換処理により取得された変換文章を、複数の段落に区分する区分処理プログラム、前記区分処理で区分されたそれぞれの前記段落を要約する要約処理プログラム、要約と音声データとの対応を取得する対応取得処理プログラム、要約を各段落の要約文毎に編集する文章編集処理プログラム、文書編集処理の編集結果に従って、前記対応に基づいて、前記音声データを編集するデータ編集処理プログラム、音声データや編集処理によって編集された音声データに基づいて音声を出力する音声出力処理プログラム、等の各種プログラムが格納されている。

【0025】文書格納部164には、仮名漢字変換プログラムにより作成された文書や、他の装置で作成されて記憶媒体駆動装置17や通信制御装置18から読み込まれた文書、及び、音声認識結果に基づいて作成された認識文書、及び認識文書を編集した後の文章等が格納される。この文書格納部164に格納される各文書の形式は特に限定されるものではなく、HTML (Hyper Text Markup Language) 形式、JIS形式、等の各種のテキスト形式の文章データが格納可能である。音声認識辞書格納部165には、音素、単音節、単語、形態素、文節等（以下形態素等）の単位での音声データと単語等との対応音声認識辞書が格納されている。本実施形態においては、形態素毎の音声パターンと形態素が対応する形態素辞書が格納されている。この音声認識辞書は、制御部11内の音声認識部において、マイク15から入力された音声の音声データのパターンや、他の装置で作成され記憶媒体駆動装置17や通信制御装置18から読み込まれた音声データのパターンから、対応する単語を探し出し、入力された音声を単語として認識する際に使用される。

【0026】編集データ格納部166には、本実施形態による編集処理を終了した要約のデータ、音声データ、画像データがそれぞれ対応した状態で格納される。

【0027】記憶媒体駆動装置17は、CPU111が外部の記憶媒体からコンピュータプログラムや文章、音声等を読み込むための駆動装置である。記憶媒体に記憶されているコンピュータプログラムには、本実施形態の編集処理装置により実行される各種処理のためのプログラム、および、そこで使用される辞書、データ等も含まれる。ここで、記憶媒体とは、コンピュータプログラム

やデータ等が記憶される記憶媒体をいい、具体的には、フロッピーディスク、ハードディスク、磁気テープ等の磁気記憶媒体、メモリチップやICカード等の半導体記憶媒体、CD-ROMやMO、PD（相変化変換型光ディスク）等の光学的に情報が読み取られる記憶媒体、紙カードや紙テープ等の用紙（および、用紙に相当する機能を持った媒体）を用いた記憶媒体、その他各種方法でコンピュータプログラム等が記憶される記憶媒体が含まれる。本実施形態の編集処理装置において使用される記憶媒体としては、主として、CD-ROMやフロッピーディスクが使用される。記憶媒体駆動装置17は、これらの各種記憶媒体からコンピュータプログラムを読み込む他に、フロッピーディスクのような書き込み可能な記憶媒体に対してRAM113や記憶装置16に格納されているデータ等を書き込むことが可能である。

【0028】本実施形態の編集処理装置では、制御部11のCPU111が、記憶媒体駆動装置17にセットされた外部の記憶媒体からコンピュータプログラムを読み込んで、記憶装置16の各部に格納（インストール）する。そして、本実施形態による編集処理等の各種処理を実行する場合、記憶装置16から該当プログラムをRAM113に読み込み、実行するようになっている。但し、記憶装置16からではなく、記憶媒体駆動装置17により外部の記憶媒体から直接RAM113に読み込んで実行することも可能である。また、編集処理装置によっては、本実施形態の編集処理プログラム等を予めROM112に記憶しておき、これをCPU111が実行するようにしてもよい。

【0029】通信制御装置18は、他のパーソナルコンピュータやワードプロセッサ等との間で各種形式のテキストデータや、Waveデータ等の音声データ、ビットマップデータ等の画像データ等、各種データの送受信を行うことができるようになっている。入出力I/F19は、音声や音楽等の出力を行うスピーカ、表示装置14に表示された文章や記憶装置16の文書格納部164に格納された文章等の印刷を行うための印刷装置、音声データに伴って画像データを取得するための撮像装置、編集後の音声データや、音声データとそれに伴う画像データを出力する再生装置、等の各種機器を接続するためのインターフェースである。入出力I/F19において接続される印刷装置としては、レーザプリンタ、ドットプリンタ、インクジェットプリンタ、ページプリンタ、感熱式プリンタ、熱転写式プリンタ、等の各種印刷装置が使用される。

【0030】音声出力装置20は、音声データを電気信号として出力する音声出力用ICと、この音声出力用ICの出力をデジタル-アナログ変換するD/Aコンバータと、このD/Aコンバータからのアナログの出力を増幅するアンプとを備えている。そしてこのアンプの出力端子にはスピーカが接続されており、音声データに対

応した音声出力されるようになっている。

【0031】次に、上述のような構成の編集処理装置による編集処理について図面を参照して説明する。

【0032】図2は、本実施形態による編集処理のメインの動作の流れを表したフローチャートである。本実施形態による編集処理においては、図6に示すように、CPU111は、まず、編集の対象となる音声データ及び画像データを、それらのデータの作成時刻（録音や録画の時刻）のデータとともに取得する（データ取得処理）（ステップ11）。音声データ及び画像データは、外部の撮像装置によって撮像及び同時録音され、撮像装置内のフラッシュメモリ等に記憶されているものを、入出力I/F19を介して取得する。または、通信制御装置18を介してインターネットやパソコン通信を介してダウンロードしたり、FDやDVD等の記憶媒体に記憶されるものを、記憶媒体駆動装置17を介して読み込んで取得する。また、入出力I/F19を介してCCDカメラ等の撮像装置から画像データを取得するとともに、マイク15から入力される音声を制御部内の入力音声管理部において音声データに変換することによって取得してもよい。取得した音声データはRAMの音声データ格納エリア1131に記録時刻とともに格納する。また画像データは、RAMの画像データ格納エリア1135に記録時刻とともに格納する。

【0033】尚、入力音声管理部による音声から音声データへの変換は、例えば、次のようにして行われる。即ち、入力音声管理部は、マイク15から入力されるアナログの音声信号をデジタル信号に変換するA/D変換器、及び、デジタル信号に変換された音声信号から入力音声のパターンを作成するパターン作成部を備えており、パターン作成部は、音声の周波数分布を分析することで、例えば、スペクトルや基本周波数の時系列情報を特徴として抽出し、そのパターンを入力音声のデータとする。また、時系列情報を統計的特徴（平均、標準偏差、相関行列など）に変換した情報を音声データとしてもよい。更に、パワースペクトル、LPC係数、ケプストラムの情報を音声データとしてもよい。

【0034】音声データおよび画像データと、これらのデータに対応する時刻データとを取得すると、次にCPU111は、ステップ11で取得した音声データについて音声認識処理を行う。図3は、音声認識処理の動作の流れを表したフローチャートである。この図3に示すように、CPU111は、音声認識処理においては、RAMに格納される音声データのパターンのうち、対応する形態素が決まっていない部分であって対応する時刻の最も早い部分から、音声認識辞書格納部165に格納されるパターンと、順次パターンマッチングにより対比し、音声データのパターンと、各形態素のパターンそれぞれに対する類似度を算出する（ステップ131）。

【0035】次に、各形態素のパターンに対する類似度

13

から、音声データのパターンと類似度の最も大きい形態素のパターンを検出する(ステップ133)。そして、音声データは、検出した最も類似度の大きいパターンの形態素であると特定し、この形態素を音声データの時刻と対応させてRAMの所定エリアに格納する(ステップ135)。続いて、CPU111は、RAMの音声データ格納エリア1131の音声データのうち、ステップ135で対応する形態素が確定していない部分があるかどうかを調べ(ステップ137)、対応する形態素が確定していない部分がある場合(ステップ137; Y)には、ステップ131にリターンし、対応する形態素が確定していない部分について、対応する時刻の最も早い部分からパターンマッチングを行い、以下RAMの所定エリアの音声データについての全て対応する形態素が特定されるまで、ステップ131からの処理を繰り返す。音声データ格納エリア1131に格納される音声データ全体について対応する形態素が確定した場合(ステップ137; N)には、確定した形態素を対応する時刻の順に並べ直し、音声認識結果として、時刻とともにRAMの所定エリアに格納する(ステップ139)。形態素を対応する時刻に並べることによって、1または複数の文が、音声認識結果として取得される。音声認識結果と時刻とをRAMの所定エリアに格納すると、CPU111は、音声認識処理を終了して、メインのルーチンへリターンする。

【0036】音声認識処理に続いて、CPU111は、変換文章取得処理を行う(ステップ15)。この変換文章取得処理においては、ステップ13の音声認識処理で取得した音声認識結果について、自然言語処理に従って解析を行い、特定された形態素が不自然な場合には、形態素を特定し直し、音声認識結果を修正する。

【0037】次に、CPU111は、変換文章取得処理において取得した文章について、段落毎に区分する区分処理を行う(ステップ17)。図4は、本実施形態における区分処理の流れを示すフローチャートである。この図4に示すように、区分処理では、CPU111は、必要に応じて修正された音声認識結果(変換文章)を所定の数の文毎に区切って小さな仮の段落(仮段落)毎にまとめる(ステップ171)。そして、仮段落毎について文書ベクトルを作成する(ステップ173)。この文書ベクトルの作成については、特開平11-45288号公報記載の技術等の、従来より公知の技術を用いることができる。

【0038】図5は、各仮段落の文書ベクトルを作成する文書ベクトル作成処理の動作の一例を表したフローチャートである。CPU111は、文書ベクトルを作成する仮段落について、形態素解析を行って、自立語、名詞句、複合名詞句等を含めた候補語(句)を抽出しRAM113の所定の作業領域に格納する(ステップ1731)。そして抽出した候補語(句)の仮段落中での出現

14

頻度、評価関数から、各候補語(句)重要度 $f(x)$ を決定する(ステップ1733)。ここで、評価関数としては、例えば、所定の重要語が予め指定されている場合にはその重要語に対する重み付け、単語、名詞句、複合名詞句等の候補語(句)の種類による重み付け等が使用される。

【0039】更にCPU111は、決定した重要度 $f(x)$ の値から仮段落中のキーワード a, b, \dots を決定する(ステップ1735)。そして、各キーワードの重要度 $f(x)$ を要素として、文書ベクトル $V = (f(a), f(b), \dots)$ をRAM113の所定作業領域に格納する(ステップ1735)。この文書ベクトル V を求める処理を全ての仮段落について行う。

【0040】全ての仮段落の文書に対して文書ベクトル V を求めた後、CPU111は、時刻が互いに連続する各仮段落同士の類似度 s を求める(ステップ175)。この仮段落どうしの類似度については、特開平11-45288号公報記載の技術等の、従来より公知の技術を用いることができる。本実施形態においては、各仮段落間の類似度 s を、両者の文書ベクトル b_n と文書ベクトル b_{n+1} 間の角度に依存するコサインにより求める。すなわち、両文書ベクトル b_n と b_{n+1} 間の角度を q とし、両文書ベクトルの内積を $b_n \cdot b_{n+1}$ とし、両文書ベクトルの大きさをそれぞれ $|b_n|$ 、 $|b_{n+1}|$ とした場合、両文書ベクトルの類似度 s は次の数式1により求まる。

【0041】

【数1】類似度 $s = \cos(q) = (b_n \cdot b_{n+1}) / (|b_n| \times |b_{n+1}|)$

【0042】この類似度 s の値は $-1 \leq s \leq 1$ までの値をとり、1に近いほど2つの仮段落の文書ベクトルが互いに平行に近く、2つの仮段落どうしは似ていると考えることができる。

【0043】次に、CPU111は、算出した類似度 s から、時刻において前後し、且つ類似度が高い仮段落どうしは類似しているとして同一の段落に纏めて、変換文章を区分する(ステップ177)。図6は、時刻が連続したAからGの7つの仮段落からなる変換文章について、仮段落の類似度と段落区分との関係を示した図である。図6においては、仮に類似度0.6以上の文書を類似しているものとする、図6(a)に示すように、A、B、C、Dの4文書は、全て類似度が0.6以上なので、1つの段落(X)となる。また、文書E、Fも類似度 s が0.9なので1つの段落(Y)となる。文書Gは、類似している文書が存在しないので、(Z)として独自に段落を構成する。

【0044】この実施の形態では、この段落の作成を類似度が0.6以上としたが他の値、例えば、0.7、0.8でもよく、これをユーザ側で設定できるようにしてもよい。変換文章を段落に区分した後は、区分処理を

終了してメインのルーチンにリターンする。

【0045】そして、変換文章を各段落に区分した後、CPU111は、変換文章の各段落について、要約を作成する要約作成処理を行う(ステップ19)。図7は、要約作成処理の動作を表したフローチャートである。CPU111は、まず要約作成処理を行う段落に含まれる自立語、名詞句、複合名詞句等を含めた候補語(句)を抽出しRAM113の所定作業領域に格納する(ステップ191)。そして、抽出した候補語(句)の各文書群中での出現頻度、評価関数等から、各候補語(句)重要度 $f(y)$ を決定する(ステップ193)。ここで、評価関数としては、例えば、所定の重要語が予め指定されている場合にはその重要語に対する重み付け、単語、名詞句、複合名詞句等の候補語(句)の種類による重み付け等が使用される。

【0046】更にCPU111は、決定した重要度 $f(y)$ 等から、各文書群に含まれる各センテンスに対する重要度 $F(z)$ を算出する(ステップ195)。そして、決定した各センテンスの重要度 $F(z)$ が上位から所定の要約比率(例えば、文書群の全文数の内の上位25%)以内に入る文、または重要度 $F(z)$ が所定ランク(例えば2位まで)に入る文を抽出する。そしてCPU111は、抽出した文を段落中での出現順に並べることによって当該段落についての要約とし、この要約を、そのもとになっている変換文章の各段落の時刻に対応して要約格納エリア1134に格納する(ステップ197)。この要約作成処理は、変換文章の全ての段落について行う。全ての段落についての要約を作成すると、CPU111は、メインのルーチンにリターンする。

【0047】各段落に対する要約の作成が終了するとCPU111は、これらの要約を時刻順に表示装置14に表示し、キーボードやマウスからの入力によって、各段落毎の要約の入れ換えや削除、複写を取得する。そして、入れ換え等に従って要約格納エリア1134の要約を格納し直す(ステップ21)。従って、対応する時刻も時刻順から、入れ換えや削除、複写の入力に従って、変更されたものとなる。

【0048】そして、入れ替わった時刻順に従って音声データ及び画像データを並べ替え(ステップ23)、記憶装置16の編集データ格納部166に、要約、音声データ、及び画像データを並べ替えた順順序とともに格納し(ステップ25)、編集処理を終了する。これらの画像データ、及び音声データは、キーボードやマウスからの出力コマンドの入力によって、編集された順に同時に出力され、記憶装置16の記憶媒体に書き込まれたり、音声出力装置20及び表示装置14から出力再生されたり、または、入出力I/F19に接続された再生装置によって出力再生される。

【0049】図8は、上述の編集処理装置を用いた編集処理におけるデータの状態を説明する説明図である。編

集処理においては、まず、図8(a)に示すように、データ取得処理において、編集の対象となる音声データ及び画像データが、それらのデータの作成時刻(録音や録画の時刻)とともに取得される。図8の例においては、音声及び映像が3つの内容I、II、IIIについてのものであり、Iの内容は時刻 $tk1b \sim tk1e$ に記録され、対応する音声データが音声データ(1)、画像データが画像データ(A)、IIの内容は時刻 $tk2b \sim tk2e$ に記録され、対応する音声データが音声データ(2)、画像データが画像データ(B)、IIIの内容は時刻 $tk3b \sim tk3e$ に記録され、対応する音声データが音声データ(3)、画像データが画像データ(C)となっているとする。そして、図8(b)に示すように、音声認識処理、変換文章取得処理によって、音声データに対応する変換文章データ(イ)、(ロ)、(ハ)が取得される。文書データは、図8(c)に示すように、区分処理によって段落毎に区分され、また要約作成処理によって、各段落毎の要約(イ')～(ハ')が作成される。

【0050】要約作成処理が終了すると、図8(c)に示すように、表示装置14に、各段落の要約が表示される。本例においては、表示装置には、各段落の要約に加えて、その段落名(場面)、及び撮影開始時刻を0とした場合の各段落の出力時間が各段落と同じ行に表示される。そして、これらのうち一行を指定し、ドラッグにより入れ換えると、図8(d)に示すように、この入れ換え結果が表示装置14に表示され、入れ換え結果に従った順番で、要約データ、音声データ及び画像データが、出力時間とともに編集データ格納部166に格納される。

【0051】このように、本実施形態では、音声データ及び画像データをそれらの作成時刻とともに取得する。そして、音声認識処理及び変換文章処理によって、取得した音声データに対応したテキストデータを作成し、テキストデータをその内容に基づいて段落に区分し、テキストデータと音声データ及び画像データとの対応をそれらの作成時刻から取得し、音声データ及び画像データを内容毎に纏めて区分した。そして、テキストデータに基づいて各区分の要約を作成し、要約を表示装置14に表示し、その順番を入れ換えや削除、複写に従って、音声データ及び画像データの順番の入れ換え、削除、複写が行われる。従って、本実施形態によると、音声データをテキストデータに変換し、テキストデータを編集することによって音声データを編集できるので、編集の位置等を容易に正確に指定することができる。本実施形態によると、テキストデータは、音声データの音声を音声認識により変換したものであるため、テキストデータによって直接音声データの内容を知ることができ、直感的に取り扱うことができる。本実施形態によれば、音声データに基づいて要約が作成され、音声データや画像データが内容

に従って区切られるため、データが多い場合であっても、内容の区切りを容易に把握し、内容毎に出力順序を入れ換えたり削除したり、複写したりすることができる。本実施形態によれば、音声データ及び画像データが内容毎に区分され、内容の要約が表示されるので、内容を容易に把握しながら、編集を行うことができる。本実施形態によれば、要約の編集結果に従って、音声データ及び画像データが編集されるので、膨大なデータについても、容易に編集することができる。

【0052】尚、上述のような編集処理装置は、2カ国語放送等を作成する場合の映像と音声の同期処理、ニュース番組やドキュメンタリー番組の作成においてニュースや取材の順序を入れ換える編集処理装置、自動音声応答システムにおける応答内容の作成や変更、映画中等のせりふを指定しての映像編集等、に使用することができる。

【0053】次に、本発明の第2の実施形態について説明する。尚、本実施形態において、図1～図8に示す第1の実施形態と同様の処理部、手段、装置、及び処理等については、同一の符号を付し、説明は省略する。

【0054】図9は、本発明の第2の実施形態による編集処理の流れを示すフローチャートである。本実施形態においては、編集処理装置は、RAM113には、重要語格納エリア、要約格納エリアは設けられていない。また、編集処理においては、データ取得処理（ステップ11）においては、音声データと、その音声データの録音時刻とを取得する。この時刻は、音声データの形態素毎に対応して取得する。そして、図9に示すように、変換文章取得処理の後は、変換文章を表示装置14に表示し、形態素毎の入れ換えや削除、複写を、キーボードやマウスからの範囲及びコマンドの選択や入力によって取得する（ステップ57）。形態素毎の入れ換えや削除、複写を取得すると、CPU111は、これらの入力に従って、対応する音声データを録音時刻から調べ、音声データを並べ替え、編集データ格納部166に格納し、編集処理を終了する。

【0055】図10は、本実施形態の編集処理装置を用いた編集処理におけるデータの状態を説明する説明図である。図10の例においては、データ取得処理（ステップ11）において、「今日私は、チョコレートを食べた。」という音声を録音した音声データを取得している。そして、図10(a)に示すように、音声認識処理（ステップ13）によって、取得した音声データのパターンに対応する形態素を取得する。図10に示す例においては、「今日は他誌は、チョコレートを食べた。」という形態素が取得されたとする。形態素を取得する場合、対応する音声データ時刻 t_1 、 t_2 、・・・ t_{12} も形態素毎に取得する。そして、変換文章取得処理においては、形態素を元になる音声データの時刻順に並べた状態で自然言語処理によって、再び形態素を取得し、図

10(b)に示すように、形態素毎に区切って、変換文章を取得する。図10の例においては、自然言語処理を加えることによって、「今日は他誌は」が「今日私は」に訂正されている。そしてこの訂正に伴う形態素の変更によって、形態素毎の対応時刻も変更されている。

【0056】そして、CPU111は、変換文章を、文節毎に区切った状態で、表示装置14に表示し、キーボードやマウスからの入力によって、文節の入れ換えや削除、複写を取得する。そして、図10(c)に示すように、表示された変換文章の文節の入れ換えに従って、音声データも順序の入れ換えまたは削除を行い、編集データ格納部166に格納する。この音声データは、キーボードやマウスからの出力コマンドの入力によって、編集された順に同時に出力され、記憶装置16の記憶媒体に書き込まれたり、音声出力装置20及び表示装置14から出力再生されたり、または、入出力I/F19に接続された再生装置によって出力再生される。

【0057】この様に、本実施形態では、音声データをその作成時刻とともに取得する。そして、音声認識処理及び変換文章処理によって、取得した音声データに対応したテキストデータを作成し、テキストデータをその内容に基づいて文節に区分し、テキストデータと音声データとの対応を文節毎にそれらの作成時刻から取得した。そして、テキストデータに基づいて文章を表示装置14に表示し、文節毎の順序の入れ換えや削除、複写を取得し、この入れ換えや削除、複写に従って、音声データの順序の入れ換え、削除、複写を行う。従って、本実施形態によると、音声データをテキストデータに変換し、テキストデータを編集することによって音声データを編集できるので、編集の位置等を容易に正確に指定することができる。本実施形態では、音声データの形態素毎に対応時刻が取得され、また、テキストデータにおいても形態素に対応して対応時刻が取得されており、対応時刻を介して、形態素単位で、音声データとテキストデータとの対応を取得することができる。そして、テキストデータを入れ換えると、この入れ換えに従って、音声データが形態素単位で並べ替えられる。従って、形態素単位で音声データを編集することができる。本実施形態によれば、テキストデータは、音声データの音声を音声認識により変換したものであるため、テキストデータによって直接音声データの内容を知ることができ、直感的に取り扱うことができる。

【0058】このような、音声データを編集する編集処理装置は、思いついた事柄をレコーダ等に吹き込み、後にひとつのまとまりのある文章に編集する場合に使用することができる。また、語学学習者や幼児が文節の組み替えによって、文法を学習する学習機として、また、文章作成ゲームのゲーム機として使用することができる。

【0059】尚、本発明は、上述の実施形態に限定されるものではなく、本発明の趣旨を逸脱しない限りにおい

て適宜変更が可能である。例えば、上述の各実施形態においては編集処理装置としてコンピュータを用いているが、コンピュータに限定されるものではなく、編集処理のための専用機でもよい。上述の各実施形態においては、音声認識処理に続いて、自然言語処理によって形態素を取得し直す変換文章取得処理を行っているが、予め自然言語処理に基づいて形態素を取得した音声認識結果を得て、変換文章取得処理を行わないようにしてもよい。また、変換文章取得処理においては、形態素の区切りをユーザが指定することによって、新たに形態素を取得し直すようにしてもよい。

【0060】上述の第1の実施形態においては、CPU 111に音声認識部が設けられ、音声の入力によって音声データを取得できるようになっているが、音声認識部を備えず、音声データを取得するものとすることもできる。上述の第1の実施形態においては、類似度が所定以上どうしであって、時刻が連続している仮段落を纏めて段落を形成しているが、時刻は考慮せず、類似度が所定以上であるものについて全て仮段落を纏めて段落としてもよい。また、段落の区分は、仮段落の類似度によって決定しなくてもよく、例えば、変換文章全体から重要語を抽出し、最初の文と2つ目の文、最初から3つ目の文まで、・・・というように、重要語の重要度を求めている、重要語のうち最も重要度の高い重要語（最重要語）が切り替わる位置を段落の変わり目として、各段落の区分を求めてもよい。上述の第1の実施形態においては、各仮段落間の類似度 s を、両者の文書ベクトル b_n と文書ベクトル b_{n+1} 間の角度に依存するコサインにより求めているが、類似度はこれに限られるものではなく、例えば、各仮段落の最も需要度の高い候補語のみについて、同一の場合には類似度を1、異なる場合には類似度を0とする等としてもよい。

【0061】上述の第1の実施形態においては、各段落毎に要約を作成し、要約を参照して各段落を入れ換えたり削除、複写する編集ができるようになっているが、各段落の変換文章そのままを表示装置14に表示して、段落毎の他、文毎に入れ換えや削除、複写をしてもよい。また、第2の実施形態を兼ね備えることによって文節、形態素毎に入れ換えや削除、複写をできるようにしてもよい。

【0062】上述の実施形態及び各変形例においては、仮名漢字変換辞書格納部161、文書格納部164、音声認識辞書格納部165、編集データ格納部166が全てプログラム格納部162と同じ記憶装置に設けられているが、例えば、編集データ格納部166はフロッピーディスクや光磁気ディスク等、記憶媒体駆動装置の記憶媒体に設ける等、これらのうちの一部または全てを他の手段に設けてもよい。

【0063】上述の実施形態及び各変形例においては、入力音声は日本語となっているが、あらゆる言語につい

て、音声データを取得し、編集処理を行うことが可能である。その場合、対象となる言語用の形態素解析アルゴリズム等を使用するといった、本発明の構成には影響のない部分を変更するだけでよい。

【0064】尚、以上の実施形態において説明した、各装置、各部、各動作、各処理等に対しては、それらを含む上位概念としての各手段（～手段）により、実施形態を構成することが可能である。例えば、「CPU 111は、ステップ11で取得した音声データについて音声認識処理を行う。」との記載に対してCPU 111とは別途、音声認識処理を行う音声認識手段を構成するようにしてもよい。同様に、その他各種動作に対して「～（動作）手段」等の上位概念で実施形態を構成するようにしてもよい。

【0065】

【発明の効果】以上説明したように、本発明によれば、音声データの編集を容易に短時間で行うことが可能である。

【図面の簡単な説明】

【図1】本発明の編集処理装置の第1の実施形態であり、本発明の編集処理プログラムが記憶された記憶媒体の第1の実施形態の該プログラムが読み取られた、コンピュータの構成を表したブロック図である。

【図2】図1の実施形態による編集処理のメインの動作の流れを表したフローチャートである。

【図3】図1の実施形態による音声認識処理の動作の流れを表したフローチャートである。

【図4】図1の実施形態における区分処理の流れを示すフローチャートである。

【図5】各仮段落の文書ベクトルを作成する文書ベクトル作成処理の動作の一例を表したフローチャートである。

【図6】時刻が連続したAからGの7つの仮段落からなる変換文章について、仮段落の類似度と段落区分との関係を示した図である。

【図7】図1の実施形態における要約作成処理の動作を表したフローチャートである。

【図8】図1の実施形態の編集処理におけるデータの状態を説明する説明図である。

【図9】本発明の第2の実施形態による編集処理の流れを示すフローチャートである。

【図10】図9の実施形態による編集処理におけるデータの状態を説明する説明図である。

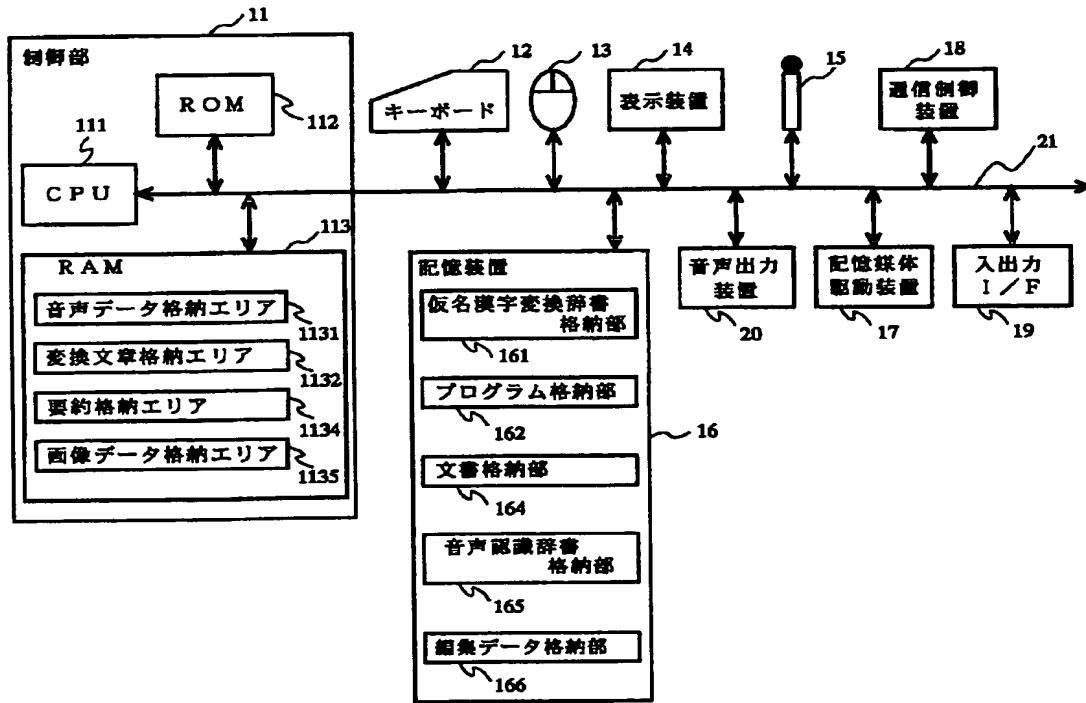
【符号の説明】

| | |
|------|------------|
| 11 | 制御部 |
| 112 | ROM |
| 113 | RAM |
| 1131 | 音声データ格納エリア |
| 1132 | 変換文章格納エリア |
| 1134 | 要約格納エリア |

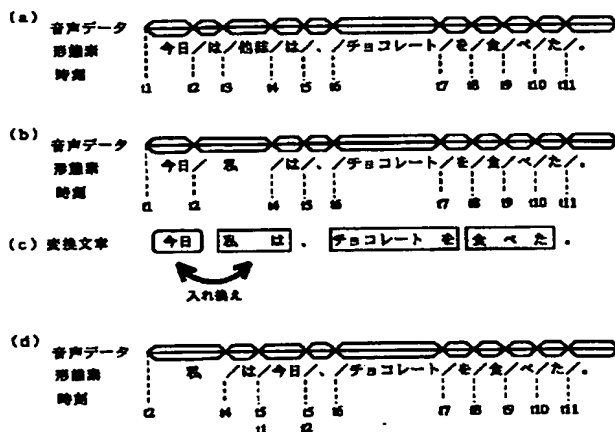
| | | | |
|---------|-------------|---------|-----------|
| 21 | | 22 | |
| 1 1 3 5 | 画像データ格納エリア | * 1 6 4 | 文書格納部 |
| 1 2 | キーボード | 1 6 5 | 音声認識辞書格納部 |
| 1 3 | マウス | 1 6 6 | 編集データ格納部 |
| 1 4 | 表示装置 | 1 7 | 記憶媒体駆動装置 |
| 1 5 | マイク | 1 8 | 通信制御装置 |
| 1 6 | 記憶装置 | 1 9 | 入出力 I / F |
| 1 6 1 | 仮名漢字変換辞書格納部 | 2 0 | 音声出力装置 |
| 1 6 2 | プログラム格納部 | * | |

*

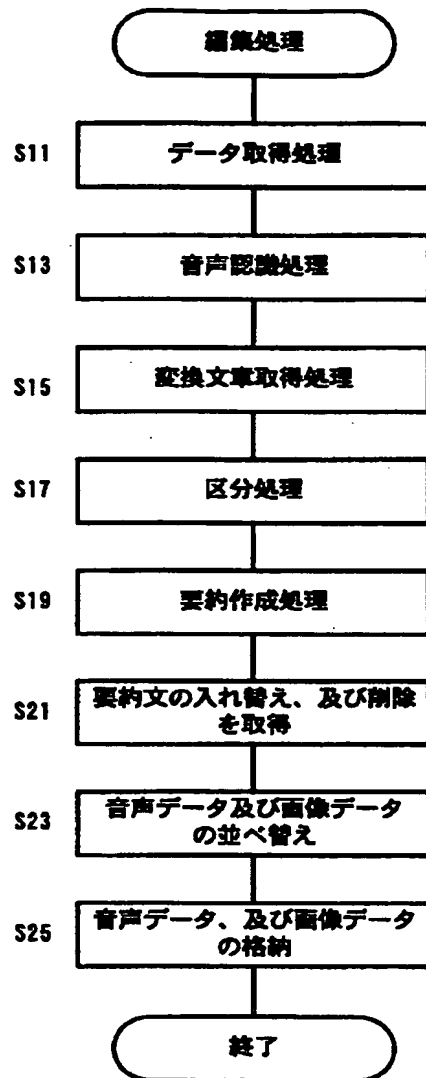
【図1】



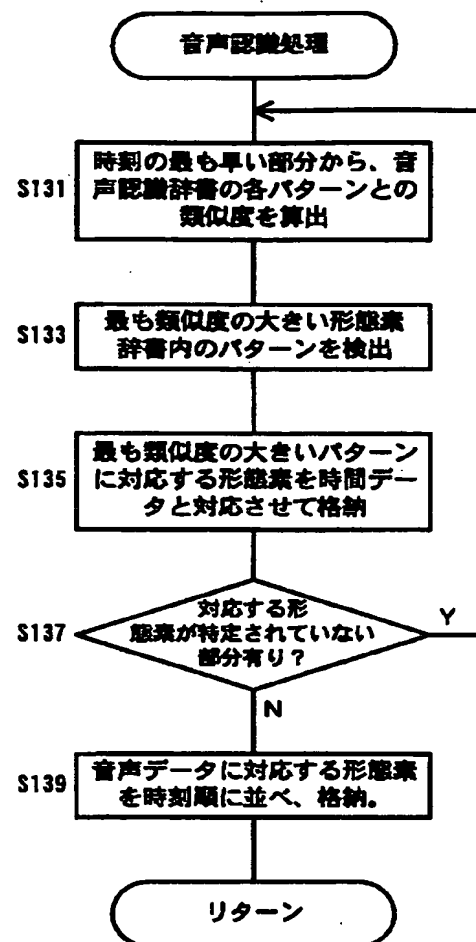
【図10】



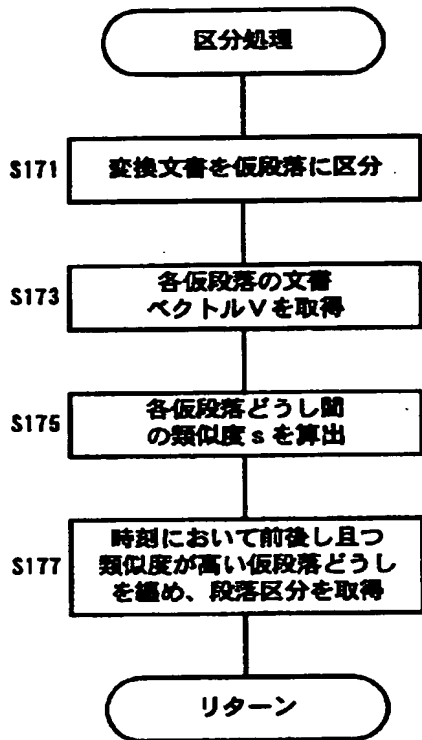
【図2】



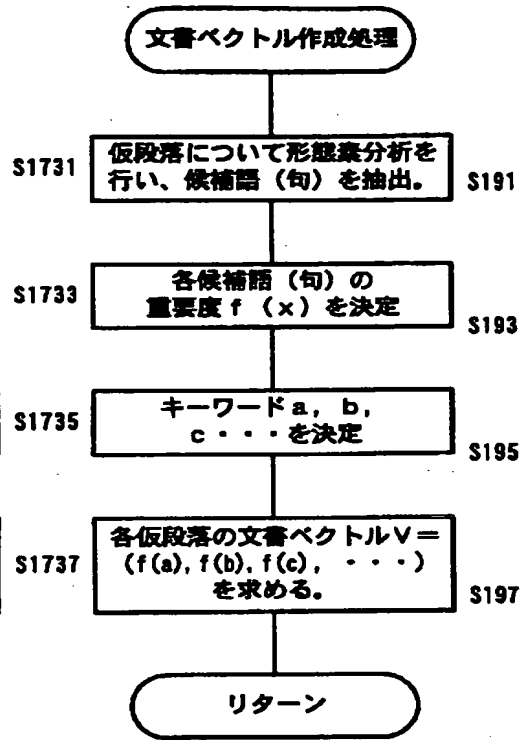
【図3】



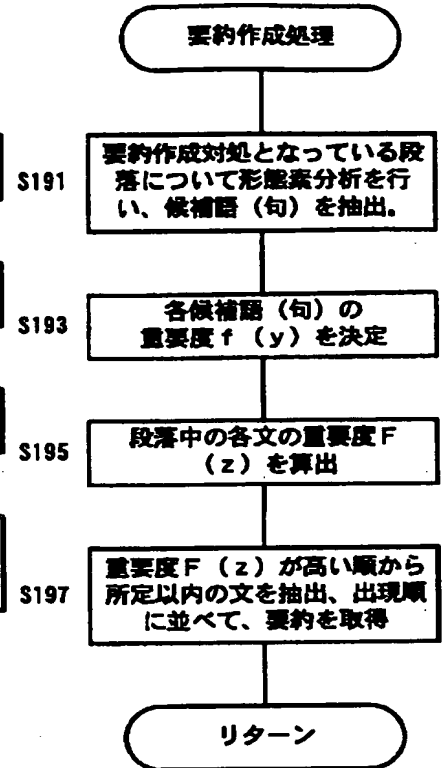
【図4】



【図5】



【図7】

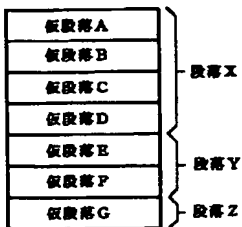


【図6】

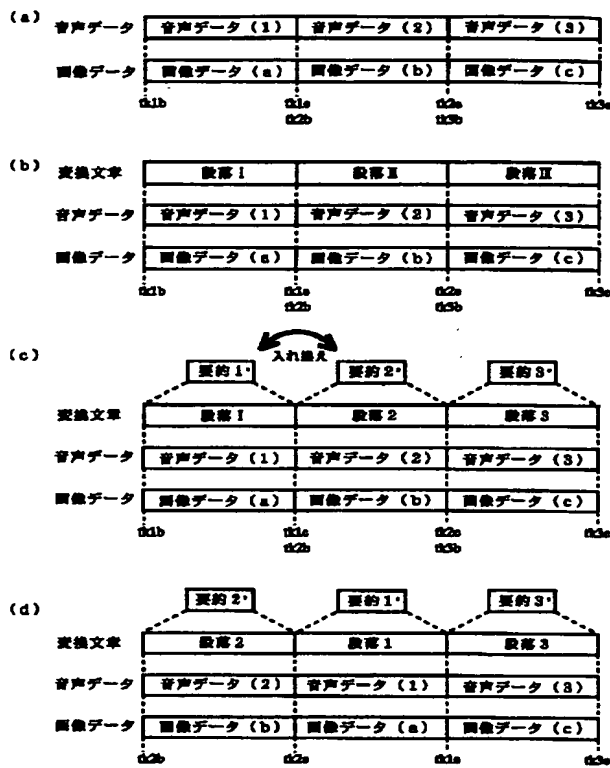
(a)

| | A | B | C | D | E | F | G |
|---|------|------|-----|------|------|------|------|
| A | | 0.9 | 0.8 | 0.8 | -0.2 | 0.1 | -0.2 |
| B | 0.9 | | 0.9 | 0.7 | -0.3 | 0.4 | 0.1 |
| C | 0.8 | 0.9 | | 0.7 | 0.1 | 0.3 | 0.1 |
| D | 0.8 | 0.7 | 0.7 | | 0.2 | -0.4 | 0.3 |
| E | -0.2 | -0.3 | 0.1 | 0.2 | | 0.9 | -0.3 |
| F | 0.1 | 0.4 | 0.3 | -0.4 | 0.9 | | -0.4 |
| G | -0.2 | 0.1 | 0.1 | 0.3 | -0.3 | -0.4 | |

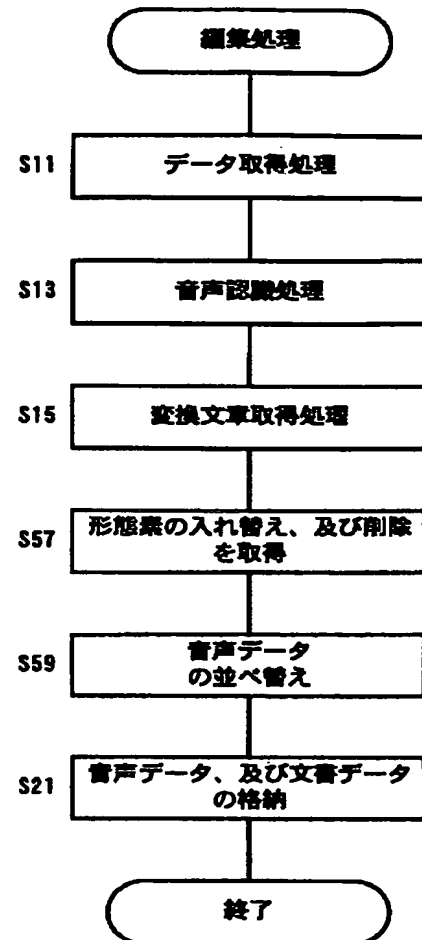
(b)



【図8】



【図9】



フロントページの続き

(72)発明者 東 宏一
徳島県徳島市沖浜東3丁目46番地 株式会
社ジャストシステム内

Fターム(参考) 5B009 KB05
5D015 KK02
9A001 BB01 BB03 BB04 BB05 CC07
DD02 DD13 EE02 EE05 FF03
GG05 HH11 HH12 HH15 HH16
HH17 HH23 HH33 JJ25 JJ29
JJ35 JJ74 JJ76 KK02 KK09
KK15 KK16 KK42 KK43